

Title of the Invention

SHEET-RECOGNIZING MACHINE

Inventors

Akihiro KAWAOKA  
Tetsuo MACHIDA  
Tetsuro KIYOMATSU  
Toshinori KAJIURA

- 1 -

## SHEET-RECOGNIZING MACHINE

## BACKGROUND OF THE INVENTION

The present invention relates to a sheet-recognizing machine and more particularly to a sheet-recognizing machine or system for recognizing the type of a sheet or slip from acquired image data in accordance with sheet definition data stored in advance.

Generally, the format of a sheet such as a bill for public utilities charges and a statement of payment for taxes is defined by and changes with each business company or each local government. There are therefore several thousands types of sheet formats in the nation.

In order to realize efficient business management, financial agencies such as banks introduce an image processing system to process sheets. A sheet image processing system prepares a database storing sheet definition data for defining the types of sheets to automatically distinguish the type of each sheet by referring to the database. For example, refer to JP-A-11-167603.

## SUMMARY OF THE INVENTION

For sheet image processing and recognition, sheet type definition data and the like are prepared

and the sheet type is distinguished by using a database which stores sheet type definition data. However, as described earlier, since there are a large number of sheet format types, it is practically difficult to  
5 store definition data of all sheet types. If the format of a sheet is changed or a sheet very similar to an already existing sheet appears, it is necessary to change the definition data. It is therefore necessary to add or modify the database.

10 It takes a labor to store or modify the definition data in the database. Initial investment in configuring the database and maintenance cost become high. Furthermore, it is important to quickly collect the formats of currently circulating sheets and reflect  
15 them upon the database so that there is no sheet which cannot be distinguished by using the database.

The present invention has been made to solve the above-described problems and provides a sheet image processing machine with reduced database configuring  
20 and maintenance costs. In order to solve the above-described problems, the following means are adopted.

A sheet recognizing system has: a sheet definition managing apparatus including a sheet definition data generating apparatus for generating  
25 sheet definition data representative of a relation between a data writing field of a sheet and the written data and a charge managing apparatus for claiming a charge in accordance with a use degree of the sheet

definition data by a sheet recognizing and processing apparatus; a communication network for transmitting the sheet definition data managed by the sheet definition data managing apparatus to the sheet recognizing and  
5 processing apparatus; and a sheet recognizing apparatus for recognizing the type of the sheet by referring to the sheet definition data acquired via the communication network. The sheet image processing apparatus stores image data of a sheet whose type  
10 cannot be distinguished by using the acquired information, and can transmit the stored image data to the sheet definition data managing apparatus.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a diagram showing a sheet image  
15 processing system according to an embodiment of the invention.

Figs. 2A and 2B are diagrams showing examples of a sheet to be processed.

Figs. 3A and 3B are diagrams showing data  
20 which is representative of characteristic fields and necessary for distinguishing between sheets.

Figs. 4A and 4B are diagrams showing data for defining characteristic fields.

Figs. 5A and 5B are diagrams showing fields  
25 required by application programs.

Figs. 6A and 6B are diagrams showing data which defines fields required by application programs.

Fig. 7 is a diagram showing a character recognizing and confirming display screen.

Fig. 8 is a diagram illustrating a sheet recognizing process to be executed by a sheet image  
5 processing system.

Fig. 9 is a flow chart illustrating a process to be executed by a sheet image managing server.

Fig. 10 is a flow chart illustrating an operation of a sheet defining terminal.

10 Fig. 11 is a flow chart illustrating a process to be executed by a sheet definition data verifying terminal.

Fig. 12 is a diagram showing a sheet management table.

15 Fig. 13 is a diagram showing a charge table.

Figs. 14A and 14B are diagrams showing management tables.

Fig. 15 is a flow chart illustrating a process to be executed by a charge information managing  
20 server.

Fig. 16 is a flow chart illustrating a process to be executed by a recognizing server.

#### DESCRIPTION OF THE EMBODIMENTS

An embodiment of the invention will be  
25 described with reference to Figs. 1 to 16. Fig. 1 is a diagram showing a sheet image processing apparatus or sheet image processing system according to the

embodiment of the invention. In Fig. 1, a sheet definition data managing apparatus 100 generates sheet definition data representative of a relation between a sheet writing position and written data, and manages  
5 the generated sheet definition data. A sheet recognizing machine or apparatus 200 acquires the sheet definition data from the sheet definition data managing apparatus 100, and distinguishes the type of a sheet read with an image scanner or the like in accordance  
10 with the acquired sheet definition data.

An operator at a sheet defining terminal 101 enters image data of a sheet or sheet definition data while looking at the sheet. A sheet definition data verifying terminal 102 verifies whether the input sheet  
15 definition data is correct or not. A sheet image managing server 103 stores sheet image data in a sheet image database 104 and manages it. The sheet image data is read with an image scanner 105 or acquired from the sheet recognizing apparatus 200 via networks 10 and  
20 11. A sheet image database 104 stores sheet image data, and the scanner 105 reads the image of a sheet. A sheet definition data managing server 106 stores the sheet definition data input from the sheet defining terminal 101 in a sheet definition database (master)  
25 107 and manages the stored data. The data stored in the sheet definition database 107 is edited to classify it into the data to be used for each sheet recognizing apparatus 200 and supply it the classified data to the

apparatus 200. Although only one sheet recognizing apparatus 200 is shown in Fig. 1, a plurality of sheet recognizing apparatus 200 are provided in practice. The data stored in the sheet definition database 107 is therefore edited to classify it into the data to be used for each sheet recognizing apparatus 200. A charge information managing server 111 manages charge data of each sheet recognizing apparatus 200 in accordance with the number of use times of the sheet definition data.

A recognizing server 201 distinguishes the type of a sheet read with an image scanner 204 in accordance with the sheet definition data acquired from the sheet definition data managing server 100. A sheet definition database (individual database) 202 provided for each sheet recognizing apparatus 200 stores the sheet definition data for the apparatus 200 acquired from the sheet definition data managing server 100. A sheet image processing terminal 203 transmits image data of a sheet read with the image scanner 204 to the recognizing server 201 via the network 12. The sheet image processing terminal 203 has an application program which provides various services such as a municipal tax paying process. A unrecognizable image file 210 stores image data of a sheet which the sheet recognizing server 201 cannot distinguish even if the sheet definition data is utilized. The sheet recognizing server 201 manages the sheet definition

database 202 and unrecognizable image file 210. The recognizing server 201 also counts the number of use times of the sheet definition data stored in the sheet definition database 202 and transmits the number of use  
5 times to a charge information managing server 111.

Figs. 2A and 2B are diagrams showing examples of a sheet to be processed by the sheet image processing apparatus. Fig. 2A shows a OO city individual municipal tax payment sheet 300 and Fig. 2B  
10 shows a AA city individual municipal tax payment sheet 350. Each sheet is made of three separable sheets "Receipt Certificate", "Payment Sheet" and "Paid Notice". As shown in Figs. 2A and 2B, although the sheets of two cities resemble closely, the city names  
15 are different and the positions and the like of a municipal code frame, an account number frame, a designation number frame, and a subscriber name frame are also different. This is because OO city forms a payment sheet specific to OO city, and AA city forms a  
20 payment sheet specific to AA city.

Figs. 3A and 3B are diagrams showing the data (characteristic field data) which is representative of characteristic fields and necessary for distinguishing between sheets shown in Figs. 2A and 2B, and Figs. 4A  
25 and 4B are diagrams showing the data for defining the characteristic fields shown in Figs. 3A and 3B.

The sheet recognizing apparatus 200 distinguishes the type of a sheet, e.g., a sheet ID,



from the characteristic field data shown in Figs. 3A and 3B by referring to the layout definition data shown in Figs. 4A and 4B. Layout definition data 410 shown in Fig. 4A defines the characteristic field (Fig. 3A) in the OO city individual municipal tax payment sheet. Of the layout definition data, "sheet ID" is a unique number assigned to the "OO city individual municipal tax payment sheet". "Size" defines the size of the sheet. In this example, the unit is 1/10 mm.

- 10 "Character 1" represents characters "OO city individual municipal tax" 310 shown in Fig. 3A and defines a distance from an origin (upper left) and a character string. "Rectangle 1" represents a rectangle 311 shown in Fig. 3A and defines a start position and an end position as distances from the origin.
- 15 "Rectangle 2" represents a rectangle 312 shown in Fig. 3A and defines a start position and an end position as distances from the origin. "Continuous rectangle 1" represents a rectangle 313 shown in Fig. 3A and defines a start position as a distance from the origin, and the number of continuous rectangles and the size of one rectangle. Layout definition data 460 shown in Fig. 2B defines the characteristic field shown in Fig. 3B in the OO city individual municipal tax payment sheet, and
- 20 is structured in a manner similar to the layout definition data for defining the characteristic fields in the OO city individual municipal tax payment sheet. By using the layout definition data shown in Fig. 3B,

the sheet image processing terminal 203 distinguishes the type of a sheet read with the scanner or the like.

Figs. 5A and 5B show fields of the sheets shown in Figs. 2A and 2B. The field is necessary for a  
5 recognizing application program to recognize the sheet. In the examples shown in Figs. 5A and 5B, Fig. 5A shows fields 320 to 323 in the OO city individual municipal tax payment sheet required by the application program, and Fig. 5B shows fields 361 to 363 in the ΔΔ city  
10 individual municipal tax payment sheet required by the application program. The application program runs on the sheet image processing terminal 203 and provides services such as a municipal tax paying process.

Figs. 6A and 6B are diagrams showing data  
15 (format definition data) for defining the fields shown in Figs. 5A and 5B required by the application program. The sheet recognizing apparatus 200 recognizes the information written in each field of a sheet by referring to the format definition data. In Fig. 6A,  
20 reference numeral 420 represents format definition data of the OO city individual municipal tax payment sheet shown in fig. 2A. In Fig. 6A, "Field 1" represents the "account number" 320 shown in Fig. 5A and defines a field attribute, a frame type, a start position, a  
25 frame size, and a font type. "Field 2" represents the "designation number" 321 shown in Fig. 5A and defines a field attribute, a frame type, a start position, a frame size, and a font type. "Field 3" represents the

"amount of money" 322 shown in Fig. 5A and defines a field attribute, a frame type, a start position, a frame size, a font type and the number of characters.

"Field 4" represents the "address name" 323 shown in

5 Fig. 5A and defines a field attribute, a frame type, a start position, a frame size, a font type and the start and end positions of pre-print to be deleted.

Reference numeral 470 represents the format definition data of the ΔΔ city individual municipal tax payment

10 sheet shown in Fig. 2B, and is structured in a manner similar to the format definition data of the OO city individual municipal tax payment sheet. Numerical values for the positions and the like are different as described with reference to Figs. 2A and 2B.

15 It is desired that the format definition data for all fields are prepared in order to allow various application programs can share the format definition data. A plurality set of format definition data may be prepared for one sheet. Although the layout definition  
20 data and format definition data are supplied separately, they may be supplied integrally. By using the format definition data shown in Figs. 6A and 6B, characters, numerals and the like in the sheets shown in Figs. 2A and 2B are recognized and the recognition  
25 results are displayed (which will be described later).

Fig. 7 is a diagram showing a character recognizing and confirming display screen of the sheet image processing terminal 203. An operator corrects

the character recognition results of the image scanner while displaying this screen on the sheet image processing terminal 203. In Fig. 7, in the upper area (painted black), cut-out image data is displayed, and in the lower area, the recognition results are displayed. The operator compares the data in both the areas and can correct the data if necessary. "?" in the "designation No" field indicates an unrecognizable character.

Fig. 8 is a diagram illustrating a sheet recognizing process to be executed by the sheet image processing apparatus or sheet recognizing system. In Fig. 8, numerals surrounded with circles such as 1 and 2 with circles (①, ②, etc.) represent an example of Steps of the sheet recognizing process. At Step 1, the sheet definition data shown in Figs. 4A and 4B and Figs. 6A and 6B is generated to configure the sheet definition database 107. At Step 2, a use contract of the sheet definition database (e.g., charge per use of the sheet definition data: 10 Yen/one use) is made between the sheet recognizing apparatus 200 side and the sheet definition data managing apparatus 100 side. At Step 3, the sheet definition data managing apparatus 100 configures a sheet definition database (individual database) storing sheet definition data satisfying the requirements of the sheet recognizing apparatus 200 side and transmits the sheet definition database to the sheet recognizing apparatus 200. At Step 4, by using

the received sheet definition database (individual database), the sheet recognizing apparatus 200 performs image processing of the image data of a sheet read with the image scanner 204 to distinguish the type of the sheet. In this case, a charge process is performed in accordance with the number of use times of the sheet definition data. The image data of a sheet which cannot be recognized by using the sheet definition database (individual database) is stored in the  
10 unrecognizable image file 210. At Step 5, the sheet definition data managing apparatus 100 acquires the image file stored in the unrecognizable image file 210, and generates sheet definition data in accordance with the image file to supplement or update the sheet  
15 definition database. At Step 6, the sheet definition data managing apparatus 100 acquires and sums up the charge information from the sheet recognizing apparatus 200. At Step 7, in accordance with the summed-up charge information, a use charge is claimed to the  
20 sheet recognizing apparatus 200 side. At Step 8, the sheet recognizing apparatus 200 side pays the claimed charge.

Fig. 9 is a flow chart illustrating the function of the sheet image managing server 103.  
25 First, at Step 501, image data of a sheet is stored in a temporary file, the image data being acquired by the image scanner 105 or acquired from the unrecognizable image file 210 via the network 10. At Step 502, the

image data is read from the temporary file and displayed on the display screen of the sheet image management server 103. At Step 503, by using an input device of the sheet image managing server 103, sheets  
5 are classified into categories as viewed from various viewpoints, by using information necessary for the sheet recognizing process such as taxes, public utilities charges, and municipal names, and enters information such as the name of a category, a sheet  
10 name, a sheet acquisition date and acquisition source information. At Step 504, such information necessary for the sheet recognizing process together with the image data is stored in the sheet image database 104. Step 501, and Steps 502 to 504 may be executed in  
15 parallel as independent processes.

Fig. 10 is a flow chart illustrating the process to be executed by the sheet defining terminal 101. At Step 511, sheet image data is read from the sheet image database 104. At Step 512, it is checked  
20 whether an instruction of an operator is new generation of sheet definition data or modification thereof. In the case of the modification, the flow advances to Step 513, whereas in the case of the new generation, the flow advances to Step 514. At Step 513, the definition  
25 data corresponding to the read image data is read from the sheet definition database 107. At Step 514, the sheet image data is read by utilizing line segment recognition techniques or the like, and by referring to

the read sheet image data, a portion of the sheet definition data is automatically generated. At Step 515, the sheet image data and generated or read sheet definition data are displayed. At Step 516, the sheet  
5 definition data is input or edited by using the input device of the sheet defining terminal. At Step 517 the sheet definition data is stored in the sheet definition database 107. The judgement at Step 512 may be executed first, and if the operator instruction is a  
10 modification instruction, the process at Step 513 is executed before the process at Step 511.

Fig. 11 is a flow chart illustrating the process to be executed by the sheet definition data verifying terminal 102. First, at Step 531, sheet  
15 definition data to be verified is read from the sheet definition database 107. At Step 532, image data is sequentially read from the sheet image database 104 to execute a sheet recognizing process by using the sheet definition data. At Step 533, the obtained process  
20 results are displayed. Whether the process results are correct or not can be judged by the operator. Whether the process results are correct or not may be automatically judged by using a program which uses stored correct data to compare sequentially read sheet  
25 image data with the correct data. If it is judged that the process results are not correct, the sheet definition data input process shown in Fig. 10 is performed again.

Fig. 12 is a diagram showing a sheet management table stored in the sheet definition data managing server 106. The sheet definition data managing server 106 has a sheet management table 450 shown in Fig. 12 which table stores sheet definition data classified into each category. By referring to this sheet management table 450, the sheet definition data managing server 106 determines that the sheet definition data belonging to which category is supplied to each sheet recognizing apparatus 200. In Fig. 12, a O symbol indicates that the data is supplied being charged, a Δ symbol indicates that the data is supplied free, and no symbol indicates that the data is still not supplied. An initial registration date of sheet definition data and data necessary for management such as data change and history may also be stored in this table.

Fig. 13 is a diagram showing a charge table. The charge table 470 stores sheet definition data represented by the sheet ID and the number of use times of each set of the sheet definition data. In addition to the number of use times, the amount of use charge may also be stored.

Figs. 14A and 14B are diagrams showing the management tables of the charge information managing server 111. A summing-up charge information table 610 shown in Fig. 14A stores the number of use times by each sheet recognizing apparatus, and a summed-up



charge information table 611 shown in Fig. 14B is used for managing the claimed charge amount of each sheet recognizing apparatus. The summing-up charge information table 610 and summed-up charge information table 611 may be formed in unison. These tables may also be integrally formed with the management table of the sheet definition data managing server 106. By referring to the charge table shown in Fig. 13 and the information tables shown in Figs. 14A and 14B, it is possible to recognize the number of use times and the use charge of the sheet definition data for each sheet recognizing apparatus so that a process to be described later can be executed such as a process of determining when the charge was claimed and whether the payment was settled.

Fig. 15 is a flow chart illustrating the process to be executed by the charge information managing server 111. First, at Step 551, at the predetermined date and time (e.g., in the first day of each month at A.M. 0), the charge information managing server 111 refers to the charge table 470 shown in Fig. 13 of the sheet recognizing apparatus 200 via the network 10, acquires data of the number of use times stored in the charge table, and writes the acquired data in the summing-up change information table 610. The sheet recognizing apparatus 200 may transmit the data of the number of use times stored in the charge table to the charge information managing server 111.

At Step 552, by referring to the summing-up charge information table 610 storing the data of the number of use times and the sheet definition database 107, the charge information managing server 111 calculates the use charge of each sheet recognizing apparatus and stores the calculation results in the summed-up charge information table 611. At Step 553, the use charge is claimed to the sheet recognizing apparatus 200, for example, via the network 10. The sheet recognizing apparatus 200 side pays the use charge and the sheet definition data managing apparatus 100 side registers the payment date in the table shown in Fig. 14 as a payment flag.

Fig. 16 is a flow chart illustrating the process to be executed by the recognizing server 201. First at Step 541, sheet image data is acquired from the sheet image processing terminal 203. At Step 542, sheet definition data is sequentially read from the sheet definition database (individual) 202, and by referring to the sheet layout information in the sheet definition data, a recognizing process is performed for the sheet read from the sheet image processing terminal 203. At Step 543, it is judged whether the sheet has the recognizable layout definition data and format definition data. If it has the recognizable data, the flow advances to Step 544, whereas if not, the flow skips to Step 546. At Step 544, by referring to the sheet format definition data in the sheet definition

data, an image of each field is cut out to perform a sheet recognizing process. At Step 545, the number of use times of the sheet definition data is incremented by "1" and stored in the charge table 470 shown in Fig.

5 13. At Step 546, the unrecognizable sheet image data is stored in the unrecognizable image file 210. At Step 547, the recognition results of the sheet image are transmitted to the sheet image processing terminal 203.

10 Although it is judged at Step 543 whether the sheet has the recognizable layout definition data and format definition data, whether the sheet has the recognizable layout definition data and whether the sheet has the recognizable format definition data may  
15 be performed at different Steps, respectively.

In the foregoing description, the sheet definition data is transmitted from the sheet definition data managing apparatus 100 to the sheet recognizing apparatus 200, and the recognizing server  
20 on the side of the sheet recognizing apparatus 200 executes a process of recognizing the type of a sheet. The recognizing server may be provided on the side of the sheet definition data managing apparatus 100. Namely, the type of a sheet is recognized by receiving  
25 the sheet image data to be recognized from the sheet recognizing apparatus without transmitting the sheet definition data to the sheet recognizing apparatus 200. In this case, the sheet recognizing apparatus 200 side

can be realized by a simple structure having only the sheet image processing terminal 203 and image scanner 204.

The sheet definition managing apparatus 100 and sheet recognizing apparatus 200 may be realized by using a single computer. The sheet recognizing program on the sheet recognizing apparatus 200 side is required to be updated depending upon the system configuration order, for example, if the sheet definition data managing apparatus 100 having the database of the embodiment is required to be operated under the conditions that the sheet recognizing apparatus 200 is already operating by using another database. In this case, if it is difficult to update the program, a data format conversion function is built in the sheet definition data managing apparatus 100 or sheet recognizing apparatus 200.

As described above, according to the embodiment, for example, the sheet definition data managing apparatus can supply the sheet recognizing apparatus with new or modified sheet definition data as latest information, and the sheet recognizing apparatus recognizes the type of a read sheet in accordance with the supplied sheet definition data. Accordingly, the initial configuration cost of the sheet definition database can be reduced and use of the database can be promoted. The recognizing server collects unrecognizable sheet image data and transmits it to the

sheet definition database to reflect it upon the database. It is therefore easy to make the sheet definition database full of latest information. The sheet definition data managing apparatus side can claim  
5 a charge for the service of supplying the sheet definition data, whereas the recognizing side can use latest sheet definition data by paying a charge corresponding to the number of use times.

As described above, according to the  
10 invention, the sheet definition data stored in the sheet definition data managing apparatus is supplied to the sheet recognizing apparatus when necessary. It is therefore possible to reduce the initial configuration cost of the sheet definition database and promote use  
15 of the database.